

WHAT IS CLAIMED IS:

1. A method for communicating data from a first compute node of a computer system comprising multiple compute nodes interconnected by an inter-node communication network to a
5 second one of the multiple compute nodes, the method comprising:
 placing the data on a full-duplex packetized interconnect
 directly connecting a CPU of the first compute node to a network
 interface connected to the inter-node communication network;
 receiving the data at the network interface; and,
10 transmitting the data to a network interface of the second
 compute node by way of the inter-node communication network.
2. A method according to claim 1 wherein the network interface and
the CPU are the only devices configured to place data on the
15 packetized interconnect.
3. A method according to claim 1 comprising transmitting the data
from the network interface to the second computer node by way of
a full-duplex communication link of the inter-node communication
20 network.
4. A method according to claim 3 comprising passing the data
through a buffer at the network interface before transmitting the
data.
25
5. A method according to claim 1 comprising, at the network
interface, determining a size of the data and, based upon the size of

the data, selecting among two or more protocols for transmitting the data.

- 5 6. A method according to claim 5 wherein the two or more protocols comprise an eager protocol and a rendezvous protocol.
7. A method according to claim 6 comprising, upon selecting the rendezvous protocol, automatically generating a Ready To Send message at the network interface of the first compute node.
- 10 8. A method according to claim 1 wherein the data comprises a raw ethernet datagram and transmitting the data comprises encapsulating the raw ethernet datagram within one or more link layer packet headers.
- 15 9. A method according to claim 8 wherein the link layer packet headers comprise InfiniBand™ link layer packet headers.
- 20 10. A method according to claim 1 wherein the data comprises a raw internet protocol datagram and transmitting the data comprises encapsulating the internet protocol datagram within one or more link layer packet headers.
- 25 11. A compute node for use in a multi-compute-node computer system; the compute node comprising:
 - a CPU;
 - a network interface; and,

a dedicated full-duplex packetized interconnect directly coupling the CPU to the network interface.

12. A compute node according to claim 11 wherein the dedicated
5 packetized full-duplex interconnect is not shared by any devices
other than the CPU and the network interface.
13. A compute node according to claim 11 comprising a memory, and
10 a facility configured to allocate eager protocol buffers in the
memory and to automatically signal to one or more other compute
nodes that the eager protocol buffers have been allocated.
14. A compute node according to claim 13 comprising a facility
15 configured to automatically associate memory protection keys with
the eager protocol buffers and a facility configured to verify
memory protection keys in incoming eager protocol messages
before writing the incoming eager protocol messages to the eager
protocol buffers.
- 20 15. A compute node according to claim 11 wherein the network
interface comprises a hardware facility at the interface configured
to encapsulate data received on the packetized interconnect in link
layer packet headers.
- 25 16. A compute node according to claim 11 wherein the network
interface comprises a buffer connected to buffer outgoing data.

17. A compute node according to claim 11 comprising a plurality of CPUs each connected to the interface by a separate dedicated full-duplex packetized interconnect.
- 5 18. A compute node according to claim 11 wherein the CPU is connected to each of a plurality of network interfaces by a plurality of dedicated full-duplex packetized interconnects.
- 10 19. A compute node according to claim 11 wherein the network interface comprises a facility configured to determine a size of data to be transmitted to another compute node and, based upon the size, to select among two or more protocols for transmitting the data to the other compute node.
- 15 20. A computer system comprising a plurality of compute nodes according to claim 11 interconnected by an inter-node data communication network, the inter-node data communication network providing at least one full-duplex data link to the network interface of each of the nodes.